

# CONTINUOUS VERSUS DISCRETE STRUCTURES I – DISCRETE EMBEDDINGS AND ORDINARY DIFFERENTIAL EQUATIONS

by

Jacky Cresson<sup>1,2</sup> and Frédéric Pierret<sup>1</sup>

**Abstract.** — We define an abstract framework called *discrete finite differences embedding* which can be used to obtain discrete analogue of formal functional relations in the spirit of category theory. For ordinary differential equations we exhibit three main discrete associate : the differential, integral or variational discrete embeddings which corresponds to classical numerical scheme including variational integrators.

<sup>1</sup> SYRTE, Observatoire de Paris, 77 avenue Denfert-Rochereau, 75014 Paris, France

<sup>2</sup> Laboratoire de Mathématiques Appliquées de Pau, Université de Pau et des Pays de l'Adour, avenue de l'Université, BP 1155, 64013 Pau Cedex, France

## Contents

1. Introduction.....	2
<b>Part I. Discrete versus continuous objects.....</b>	<b>3</b>
2. Notations.....	3
2.1. Functional spaces.....	4
3. Continuous versus discrete functions.....	4
3.1. Discrete functions and discretisation.....	4
3.2. Interpolation mappings and lifting of discrete functions.....	5
3.2.1. Interpolation mappings.....	5
3.2.2. Properties of interpolation mappings.....	6
3.3. Algebraic structures.....	6
3.3.1. Product.....	6
3.3.2. Duality mappings.....	6
4. Discrete derivatives.....	7
4.1. Formal construction.....	7
4.2. Explicit form.....	7
4.3. Properties of discrete derivatives.....	7
4.3.1. Duality.....	8
4.3.2. Kernel of discrete derivatives.....	8
4.3.3. Discrete Leibniz formula.....	8
5. Discrete antiderivatives.....	8
5.1. Discrete antiderivatives.....	9
5.2. Explicit form.....	9
5.3. Properties of discrete antiderivatives.....	9
5.3.1. Duality.....	10
5.3.2. Discrete integration by part.....	10
5.3.3. Scalar product.....	10
6. Discrete classical results.....	11
6.1. Discrete fundamental theorem of differential calculus.....	11
6.2. Discrete Dubois-Raymond lemma.....	11
<b>Part II. Discrete embedding formalisms.....</b>	<b>12</b>
7. Abstract discrete embedding.....	12

**Key words and phrases.** — Discrete embedding formalisms, numerical methods.

8. Application to ordinary differential equations : the three forms.....	13
8.1. Discrete differential embedding.....	13
8.2. Discrete integral embedding.....	14
8.3. Discrete variational embedding.....	14
8.3.1. Lagrangian systems.....	14
8.3.2. Discrete Lagrangian functionals.....	15
8.3.3. Comparison with Marsden-West definition.....	15
8.3.4. Discrete calculus of variations and discrete Euler-Lagrange equations.....	15
8.3.5. Comparison with Marsden-West definition.....	16
8.3.6. Proof of Theorem 31.....	16
<b>Part III. Higher order discrete embeddings.....</b>	<b>16</b>
9. Quadratic embedding : Trapezoidal and Midpoint methods.....	17
9.1. Quadratic discrete embedding.....	17
9.1.1. Quadratic interpolation mapping.....	17
9.1.2. Quadratic discrete derivative and integral.....	17
9.2. Discrete differential and integral embedding.....	18
10. Cubic discrete embedding : the forward Simpson numerical scheme.....	19
10.1. Cubic discrete embedding.....	19
10.1.1. Cubic interpolation mapping.....	19
10.1.2. Cubic discrete derivative and integral.....	19
10.2. Discrete differential and integral embedding.....	20
11. Conclusion and perspectives.....	21
A. Proof of Theorem 32.....	21
B. Proof of Theorem 34.....	22
References.....	23

## 1. Introduction

Many different fields of Mathematics and Applied Mathematics deal with the construction of a *discrete* analogue of a continuous notion. The aim of this article is to provide an abstract framework to deal with discrete version of ordinary differential equations. In an informal way, we are looking for a "*functor*" from the set **ODE** of ordinary differential equations to the set **FDE** of finite differences equations

$$\mathbf{ODE} \longrightarrow \mathbf{FDE}$$

in a way inspired by *category theory* [3]. It is probably possible to give a complete categorical formulation of our approach even if we only keep in mind the general strategy.

Our objective is not only theoretic. Using this abstract "arrow" that we call *discretisation* in the following, we want to precise in particular two points :

- On how many objects (sets, mappings, etc) depends a discretisation ?
- A discretisation being fixed, to obtain a better understanding of which kind of properties and structures from the continuous objects are preserved in the discrete setting.

Of course, many answers already exist in the literature. In particular, the one dealing with Numerical Analysis of differential equations and more specifically *Geometric numerical integration* [20]. However, as we will see, some discrete analogues defined by these theory are surprisingly far from being satisfying at least from the point of view of structure and formulation, even for well known results as for example discrete analogue of Lagrangian and Hamiltonian systems which is the origin of the present work. In this paper, we make a systematic comparison of our formulation with those obtained in geometric numerical integration as exposed in [20] and also some results obtained in the context of variational integrators as exposed in [25].

In order to define the "arrow" called discretisation, we follow the philosophy of *embedding formalisms* developed for example in ([8],[9],[11],[10]) by defining a general concept of *discrete embedding*. An account of this formalism for problems related to discretisation was depicted in the articles [11] and [12] using *time scale calculus*. Here, we provide a more conceptual framework, which allows us to recover many classical methods of numerical analysis in the same setting modifying only some projection and lift maps which are inherent to the definition of a discrete embedding then generalizing our previous work in this direction. Moreover, it allows us to revisit classical definitions and results in numerical analysis.

As already pointed out, some part of our approach is related to the *time-scale* calculus point of view in particular to results as exposed for example in [12] and L. Bourdin ([4],[5]). However, the time-scale calculus coincide with standard discretisation procedure of the first order in term of approximation, due to the fact that the time-scale derivatives reduces to the forward and backward Euler derivative by construction. This point is important since in numerical analysis one is interested not only to give a clear connexion between a continuous object and a discrete one but also to obtain a high order of approximation between the discrete solution and the continuous one. Our formalism naturally extends to such a setting and we give examples of order 2 and 3.

The paper is organized as follows : In the first part, we define the discrete analogue of continuous objects like functions and differential or integral operators. The main point is to give an explicit connexion between the discrete and continuous case. This is done using some particular mappings that we call discretisation and interpolation in the following. As a consequence, we are able to give without any computations on sums or classical methods of rearranging terms new formulations of classical results (discrete integration by part, discrete fundamental theorem of differential calculus, etc). In a second part, we use the previous formalism to define the discrete embedding of a formal functional. This abstract setting allows us to cover very different objects like differential equations, integral equations or Lagrangian functional. We then describe the three natural discrete way to obtain a discrete analogue of a differential equation : the differential, integral and variational case. Each of these procedure lead to different discrete realization of the same equations.

## PART I

### DISCRETE VERSUS CONTINUOUS OBJECTS

#### 2. Notations

Let  $N \in \mathbb{N}$  and  $a, b \in \mathbb{R}$  with  $a < b$  and let  $h = (b - a)/N$ . We denote by  $\mathbb{T}$  the subspace of  $\mathbb{R}$  defined by  $\mathbb{T} = h\mathbb{Z} \cap [a, b]$  where  $h\mathbb{Z} = \{hz | z \in \mathbb{R}\}$ . The elements of  $\mathbb{T}$  are denoted by  $t_k$  for  $k = 0, \dots, N$ . The indicator function of a subset  $A \subset X$  is denoted by  $1_A$  and defined by  $1_A(t) = 1$  if  $t \in A$  and 0 otherwise. We adopt the convention that a sum from an integer  $i$  to an integer strictly inferior than  $i$  is zero.

**2.1. Functional spaces.** — We denote by  $C([a, b], \mathbb{R}^d)$  the set of functions  $x : [a, b] \rightarrow \mathbb{R}^d$ ,  $d \in \mathbb{N}^*$  and by  $C^i([a, b], \mathbb{R}^d)$  the set of  $i$ -th differentiable functions.

We then consider  $C(\mathbb{T}, \mathbb{R})$  as the set of functions with value in  $\mathbb{R}$  over  $\mathbb{T}$ . We denote by  $P_n([a, b], \mathbb{R})$  the set of continuous functions which are piecewise polynomial of degree  $n$  which are defined over subintervals  $I_i^n = [t_{ni}, t_{n(i+1)}]$  such that  $\bigcup_i I_i^n = [a, b]$  and where we supposed that  $N$  is a multiple of  $n$ . We denote also by  $P_n^+([a, b], \mathbb{R})$  ( resp.  $P_n^-([a, b], \mathbb{R})$  ) the set of polynomial functions of degree  $n$  over  $[t_{ni}, t_{n(i+1)}[$  (resp.  $]t_{ni}, t_{n(i+1)}]$ ) and right (resp. left) continuous. Let  $T = T_0, \dots, T_n \in \mathbb{R}^n$  and  $F = F_0, \dots, F_n \in \mathbb{R}^n$ , we denote by  $l_i(t) = \prod_{j \neq i} \frac{(t - T_j)}{(T_i - T_j)}$ ,  $0 \leq i \leq n$ . We denote by  $P_{L,n}^F(t) = \sum_{i=0}^n F_i l_i(t)$  the Lagrange polynomial of degree  $n$  associated to  $F$  (see [14], p.21-22).

### 3. Continuous versus discrete functions

**3.1. Discrete functions and discretisation.** — Let us begin with a general definition :

**Definition 1 (Discrete function).** — A discrete function is an element  $F \in C(\mathbb{T}, \mathbb{R})$ .

By definition, a discrete function is completely characterized by the finite set  $F_k = F(t_k)$ ,  $k = 0, \dots, N$ .

In the following, we illustrate all the discrete notions with a single example given by the following discrete function :  $\mathbf{F} = \{2, 1, 3, 2, 7, 5, 2\} \in C(\mathbb{T}, \mathbb{R})$  with  $\mathbb{T} = \{0, 1, 2, 3, 4, 5, 6\}$ .

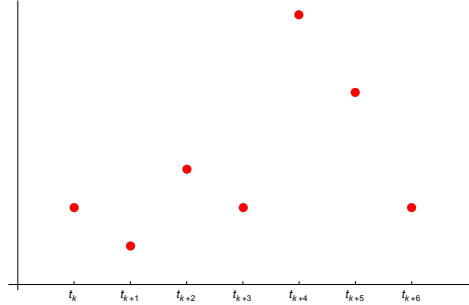


FIGURE 1. A discrete function

**Definition 2 (Discretisation of functions).** — Let  $f \in C([a, b], \mathbb{R})$ , a discretisation of  $f$  is a discrete function  $F$  such that each  $F_k$  can be computed using the value  $\{f_j = f(t_j)\}_{j=0, \dots, N}$ .

A natural or canonical discretisation is of course given by the restriction of  $f$  to  $\mathbb{T}$ .

**Definition 3 (Canonical discretisation).** — We denote by  $\pi : C([a, b], \mathbb{R}) \rightarrow C(\mathbb{T}, \mathbb{R})$  the mapping define by the restriction of a given function  $f \in C([a, b], \mathbb{R})$  to  $\mathbb{T}$ , i.e.  $\pi(f) = F$  where  $F_k = f_k$ ,  $k = 0, \dots, N$ .

The following section introduces two lift map from  $C(\mathbb{T}, \mathbb{R})$  in various finite vectorial subspaces of  $C([a, b], \mathbb{R})$ .

**3.2. Interpolation mappings and lifting of discrete functions.** — In order to construct discrete analogue of continuous mappings, we have to relate the set of discrete functions to some finite vectorial spaces of  $C([a, b], \mathbb{R})$ .

*3.2.1. Interpolation mappings.* — The notion of *interpolation* can be formally defined as follows :

**Definition 4 (Interpolation map).** — A map  $\iota : C(\mathbb{T}, \mathbb{R}) \rightarrow C([a, b], \mathbb{R})$  satisfying  $\pi \circ \iota = \text{Id}$ , where  $\text{Id}$  is the identity of  $C(\mathbb{T}, \mathbb{R})$  is called an *interpolation map*.

The name *interpolation* comes from the last condition. One can naturally extend a given interpolation mapping for discrete functions with values in  $\mathbb{R}^d$  by posing for  $F \in C(\mathbb{T}, \mathbb{R}^d)$ ,  $F = (F_1, \dots, F_d)$ ,  $F_i \in C(\mathbb{T}, \mathbb{R})$ ,  $\iota(F) = (\iota(F_1), \dots, \iota(F_d))$ .

Three classical interpolation mappings will be used in the following (see [14], Chap.II, §.1).

**Definition 5 ( $P_0^\pm$ -interpolation).** — We denote by  $\iota_0^\pm : C(\mathbb{T}, \mathbb{R}) \rightarrow P_0^\pm([a, b], \mathbb{R})$  the map defined for all  $F \in C(\mathbb{T}, \mathbb{R})$  by

$$(1) \quad \iota_0^+(F) = \sum_{i=0}^{N-1} F_i 1_{[t_i, t_{i+1}[} \quad \left( \text{resp. } \iota_0^-(F) = \sum_{i=1}^N F_i 1_{]t_{i-1}, t_i]} \right).$$

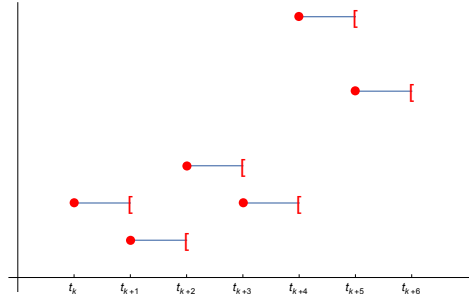


FIGURE 2.  $P_0^+$ -interpolation mapping

We need sometimes more regularity.

**Definition 6 ( $P^1$ -interpolation).** — We denote by  $\iota_1 : C(\mathbb{T}, \mathbb{R}) \rightarrow P^1([a, b], \mathbb{R})$  the map defined for all  $F \in C(\mathbb{T}, \mathbb{R})$  by

$$(2) \quad \iota_1(F)(t) = \sum_{k=0}^{N-1} \left[ F_k + \frac{F_{k+1} - F_k}{h} (t - t_k) \right] 1_{[t_k, t_{k+1}]}(t).$$

This map is of course natural and is classical in numerical analysis. The main problem is that in this case, the derivative is not defined but only left and right derivatives. This point will induce several complications in the connexion between the various discrete operators.

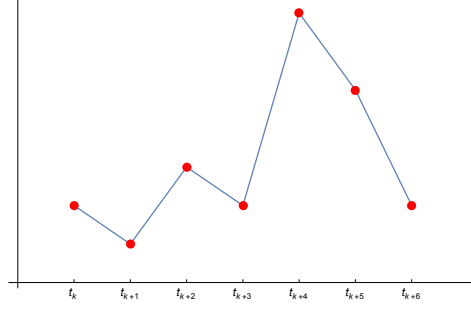


FIGURE 3. Order one interpolation mapping

**3.2.2. Properties of interpolation mappings.** — The interpolation mappings satisfy some additional interesting properties :

**Lemma 7.** — *The restriction of  $\iota_0^+ \circ \pi$  and  $\iota_0^- \circ \pi$  to  $P_+^0([a, b])$  and  $P_-^0([a, b])$  respectively is the identity.*

### 3.3. Algebraic structures. —

**3.3.1. Product.** — We define a product on  $C(\mathbb{T}, \mathbb{R})$  which is compatible with the classical product of functions. Let  $\star$  be the product defined by

$$\begin{array}{ccc}
 P_1([a, b], \mathbb{R}) \times P_1([a, b], \mathbb{R}) & \xrightarrow{\quad \cdot \quad} & P_1([a, b], \mathbb{R}) \\
 (\iota_1(F), \iota_1(G)) & & \iota_1(F) \cdot \iota_1(G) \\
 \uparrow \iota_1 & & \downarrow \pi \\
 C(\mathbb{T}, \mathbb{R}) \times C(\mathbb{T}, \mathbb{R}) & \xrightarrow{\quad \star \quad} & C(\mathbb{T}, \mathbb{R}) \\
 (F, G) & & F \star G
 \end{array}$$

**Definition 8.** — *The product  $\star$  on  $C(\mathbb{T}, \mathbb{R})$  is defined for all  $(F, G) \in C(\mathbb{T}, \mathbb{R}) \times C(\mathbb{T}, \mathbb{R})$  by  $(F \star G)_k = F_k \cdot G_k$  for  $k = 0, \dots, N$ .*

Using this definition, we can easily define the product of two vectorial discrete functions :

**Definition 9.** — *The product  $\langle \cdot, \cdot \rangle_\star$  is defined for all  $(F, G) \in C(\mathbb{T}, \mathbb{R}^d)$  by  $\langle F, G \rangle_\star = \sum_{i=1}^d F_i \star G_i$  where  $F = (F_1, \dots, F_d)$  and  $G = (G_1, \dots, G_d)$ .*

We will see that this discrete product can be used to extend the Hilbert structure to discrete functions.

**3.3.2. Duality mappings.** — Let  $\mathbb{T} \subset [a, b]$  be given,  $\mathbb{T} = \{t_0, \dots, t_N\}$ . We denote by  $\mathbb{T}_+$  and  $\mathbb{T}_-$  the subset of  $\mathbb{T}$  given by  $\mathbb{T}_+ = \{t_0, \dots, t_{N-1}\}$  (resp.  $\mathbb{T}_- = \{t_1, \dots, t_N\}$ ). We define  $C(\mathbb{T}_+, \mathbb{R})$  and  $C(\mathbb{T}_-, \mathbb{R})$ . Many properties of discrete objects come from the date of some natural mappings between these two functional spaces.

**Definition 10 (Duality mappings).** — We denote by  $\sigma : C(\mathbb{T}_+, \mathbb{R}^d) \rightarrow C(\mathbb{T}_-, \mathbb{R}^d)$  (resp.  $\rho : C(\mathbb{T}_-, \mathbb{R}^d) \rightarrow C(\mathbb{T}_+, \mathbb{R}^d)$ ) the map defined by

$$(3) \quad \sigma(F)(t_i) = F(t_{i-1}), \quad i = 1, \dots, N \quad (\text{resp. } \rho(F)(t_i) = F(t_{i+1}), \quad i = 0, \dots, N-1).$$

**Lemma 11.** — The maps  $\sigma$  and  $\rho$  are bijective and satisfy  $\rho \circ \sigma = \sigma \circ \rho = \text{Id}$ .

The proof follows from the definition of the action of  $\sigma$  and  $\rho$  on

#### 4. Discrete derivatives

**4.1. Formal construction.** — This function is not derivable but we can always define a left and right derivative in the points  $t_k$ . As a consequence, we can define  $\Delta$  and  $\nabla$  the two discrete operators corresponding to the left and right derivative denoted by  $\frac{d^+}{dt}$  and  $\frac{d^-}{dt}$  by a commutative diagram :

**Definition 12.** — The forward (resp. backward) discrete derivative  $\Delta$  (resp.  $\nabla$ ) is defined by

$$(4) \quad \Delta = \pi \circ d^+ \circ \iota_1 \quad (\text{resp. } \nabla = \pi \circ d^- \circ \iota_1).$$

This definition corresponds to the following commutative diagram

$$\begin{array}{ccc} P_1([a, b], \mathbb{R}) & \xrightarrow{\frac{d^+}{dt}} & P_0^+([a, b], \mathbb{R}) \\ \iota_1(F) \uparrow & & \downarrow \frac{d^+}{dt}[\iota_1(F)] \\ \uparrow \iota_1 & & \downarrow \pi \\ C(\mathbb{T}, \mathbb{R})_F & \xrightarrow{\Delta} & C(\mathbb{T}_+, \mathbb{R})_{\Delta F} \end{array}$$

We deduce easily from the definition that :

**Lemma 13.** — The forward (resp. backward) discrete derivative satisfies  $\Delta : C(\mathbb{T}, \mathbb{R}) \rightarrow C(\mathbb{T}_+, \mathbb{R})$  (resp.  $\nabla : C(\mathbb{T}, \mathbb{R}) \rightarrow C(\mathbb{T}_-, \mathbb{R})$ ) where  $\mathbb{T}_+ = \{t_0, \dots, t_{N-1}\}$  (resp.  $\mathbb{T}_- = \{t_1, \dots, t_N\}$ ).

*Proof.* — We have  $\frac{d^+}{dt}(P^1([a, b]) = P_+^0([a, b])$  and  $\frac{d^-}{dt}(P^1([a, b]) = P_-^0([a, b])$ . We deduce that  $\pi(P_+^0([a, b])) = C([a, b] \cap \mathbb{T}, \mathbb{R})$  and  $\pi(P_-^0([a, b])) = C([a, b] \cap \mathbb{T}, \mathbb{R})$ .  $\square$

**4.2. Explicit form.** — It is easy to obtain an explicit form for these two discrete derivatives :

**Lemma 14.** — Let  $F \in C(\mathbb{T}, \mathbb{R})$ , we have  $(\Delta F)_k = \frac{F_{k+1} - F_k}{h}$ , for  $k = 0, \dots, N-1$ , and  $(\nabla F)_k = \frac{F_k - F_{k-1}}{h}$ , for  $k = 1, \dots, N$ .

We recover the classical forward and backward derivatives used in numerical analysis.

**4.3. Properties of discrete derivatives.** —

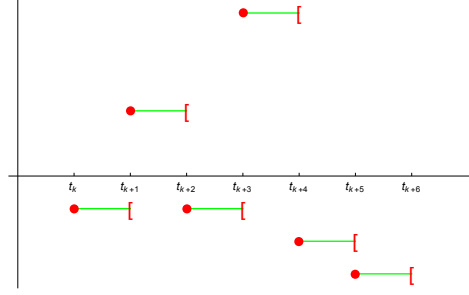


FIGURE 4. Forward discrete derivative

**4.3.1. Duality.** — The duality between  $C(\mathbb{T}_+, \mathbb{R})$  and  $C(\mathbb{T}_-, \mathbb{R})$  can be used to obtain a duality between the  $\Delta$  and  $\nabla$  derivative. Precisely, we have :

**Lemma 15.** — *Let  $F \in C(\mathbb{T}, \mathbb{R})$ , then  $\nabla(\sigma(F)) = \Delta(F)$  and  $\nabla(F) = \Delta(\rho(F))$  over  $\mathbb{T}^\pm$ .*

**4.3.2. Kernel of discrete derivatives.** —

**Lemma 16.** — *Let  $F \in C(\mathbb{T}, \mathbb{R})$ . We have  $\Delta F = 0$  (resp.  $\nabla F = 0$ ) if and only if  $F = \mathbb{F}_0$ , where  $\mathbb{F}_0$  is defined by  $\pi(F_0 Id)$  where  $Id$  is the identity function. We will also say that  $F$  is constant.*

The previous Lemma can be used for example to characterize *discrete* first integrals of classical numerical scheme (see for example ([7], Theorem 12 p.885) and compare with ([20], Theorem 6.7 p.197)).

**4.3.3. Discrete Leibniz formula.** — An important algebraic property of the classical derivative is the Leibniz formula. The following theorem gives the discrete version of this formula which mimics exactly the continuous one.

**Theorem 17 (Discrete Leibniz rule).** — *Let  $F, G \in C(\mathbb{T}, \mathbb{R})$ , we have*

$$(5) \quad \begin{aligned} \Delta(F \star G) &= (\Delta F) \star G + \sigma(F) \star (\Delta G) \text{ over } \mathbb{T}_+, \\ \nabla(F \star G) &= (\nabla F) \star G + \rho(F) \star (\Delta G) \text{ over } \mathbb{T}_-. \end{aligned}$$

It is interesting to notice that despite the constant use of the previous formula in many computations concerning numerical scheme, very few exhibit the fact that this is a discrete Leibniz formula. This is due in part to the fact that most of the computations are usually made directly using summations and not in an abstract way. This phenomenon is particularly visible in the derivation of *variational integrators* (see [25]) in the context of the *discrete calculus of variations*.

## 5. Discrete antiderivatives

Using the same procedure as for discrete derivatives, we define discrete antiderivatives.



**5.1. Discrete antiderivatives.** — Using the lift mappings we can easily define an antiderivative :

**Definition 18 (Discrete antiderivative).** — *The discrete  $\Delta$ -antiderivative (resp.  $\nabla$ -antiderivative) denoted by  $J_\Delta$  (resp.  $J_\nabla$ ) is defined by*

$$(6) \quad J_\Delta = \pi \circ \int_a^t \circ \iota_0^+ \quad \left( \text{resp. } J_\nabla = \pi \circ \int_a^t \circ \iota_0^- \right).$$

This definition for  $J_\Delta$  corresponds to the following diagram

$$\begin{array}{ccc} P_0^+([a, b], \mathbb{R}) & \xrightarrow{\int_a^t} & P_1^+([a, b], \mathbb{R}) \\ \iota_0^+(F) \uparrow & & \downarrow \pi \\ C(\mathbb{T}, \mathbb{R})_F & \xrightarrow{J_\Delta} & C(\mathbb{T}, \mathbb{R})_{J_\Delta F} \end{array}$$

$\int_a^t$  is labeled  $\int_a^t \iota_0^+(F) dt$  above the arrow.

An analogous diagram is obtained for  $J_\nabla$ .

**5.2. Explicit form.** — An explicit formula for  $J_\Delta$  (resp.  $J_\nabla$ ) is easily obtained :

**Lemma 19.** — *Let  $F \in C(\mathbb{T}, \mathbb{R})$ , we have*

$$[J_\Delta(F)]_k = \sum_{i=0}^{k-1} (t_{i+1} - t_i) F_i \quad \left( \text{resp. } [J_\nabla(F)]_k = \sum_{i=1}^k (t_i - t_{i-1}) F_i \right)$$

for  $k = 0, \dots, N$ .

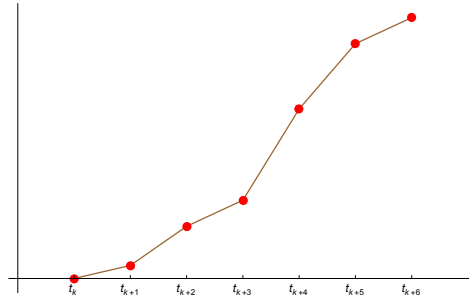


FIGURE 5. Discrete  $\Delta$ -integral

**5.3. Properties of discrete antiderivatives.** —

**5.3.1. Duality.** — The duality between the  $\Delta$  and  $\nabla$  derivatives induces a corresponding phenomenon for the  $\Delta$  and  $\nabla$  antiderivatives.

**Proposition 20.** — For all  $F \in C(\mathbb{T}, \mathbb{R})$  we have  $J_\Delta \circ \sigma(F) = J_\nabla(F)$  and  $J_\nabla \circ \rho(F) = J_\Delta(F)$ .

**5.3.2. Discrete integration by part.** — As in the continuous case, the discrete Leibniz formula for discrete derivatives induces a *discrete integration by part* formula for discrete antiderivatives.

**Theorem 21 (Discrete integration by part formula).** — Let  $F \in C(\mathbb{T}, \mathbb{R})$  and  $G \in C_0(\mathbb{T}, \mathbb{R})$ , we have

$$(7) \quad \begin{aligned} [J_\Delta(F \star \Delta(G))]_N &= -[J_\Delta(F \star (\nabla G))]_N, \\ [J_\nabla(F \star \nabla(G))]_N &= -[J_\nabla(F \star (\Delta G))]_N. \end{aligned}$$

*Proof.* — We only make the proof for  $J_\Delta$ , the case of  $J_\nabla$  being similar. Using the discrete Leibniz formula for  $\Delta$  we have  $J_\Delta[F \star \Delta G] = J_\Delta[\Delta[\rho(F) \star G] - \Delta[\rho(F)] \star G]$  over  $\mathbb{T}_+$ . Using Theorem 24 and the fact that  $G \in C_0(\mathbb{T}, \mathbb{R})$ , we have  $[J_\Delta \circ \Delta(\rho(F) \star G)]_N = \rho(F)_N \cdot G_N - \rho(F)_0 G_0 = 0$ . As  $\Delta[\rho(F)] = \nabla(F)$  over  $\mathbb{T}^\pm$  by Lemma 15, this concludes the proof.  $\square$

The previous result is never stated as below in classical Textbooks about finite differences. This is due to the fact that the operators are not usually used to write such a formula but directly on summations formulae only speaking of rearranging the terms of the sum (see ([25], p.363) for a typical example).

**5.3.3. Scalar product.** — Let  $f, g \in C([a, b], \mathbb{R}^d)$ . The classical scalar product on  $L^2$  functions denoted by  $\langle \cdot, \cdot \rangle_{L^2}$  is defined by  $\langle f, g \rangle_{L^2} = \int_a^b \langle f(t), g(t) \rangle dt$ . Using the *discrete product* and the lift mappings, we can transport the  $L^2$  scalar product over discrete functions.

**Definition 22 (Discrete scalar product).** — We call forward (resp. backward) discrete scalar product the bilinear mapping defined for all  $F, G \in C(\mathbb{T}, \mathbb{R}^d)$  by

$$(8) \quad \langle F, G \rangle_+ = J_\Delta[\langle F, G \rangle_\star]_N \quad (\text{resp.} \quad \langle F, G \rangle_- = J_\nabla[\langle F, G \rangle_\star]_N).$$

**Remark 1.** — The discrete forward (resp. backward) discrete scalar product is degenerate. Indeed, the equation  $\langle F, F \rangle_+ = 0$  (resp.  $\langle F, F \rangle_- = 0$ ) induces only

$$(9) \quad F_k = 0 \quad \text{for } k = 0, \dots, N-1, \quad (\text{resp.} \quad F_k = 0 \quad \text{for } k = 1, \dots, N).$$

Using the definition of discrete antiderivatives, we prove the following lemma :

**Lemma 23.** — The following diagram commutes

$$\begin{array}{ccc} P_0^+([a, b], \mathbb{R}^d) \times P_0^+([a, b], \mathbb{R}^d) & \xrightarrow{\langle \cdot, \cdot \rangle} & \mathbb{R} \\ (\iota_0^+(F), \iota_0^+(G)) & & \langle \iota_0^+(F), \iota_0^+(G) \rangle \\ \uparrow \iota_0^+ & & \downarrow \text{Id} \\ C(\mathbb{T}, \mathbb{R}^d) \times C(\mathbb{T}, \mathbb{R}^d) & \xrightarrow{\langle \cdot, \cdot \rangle_\Delta} & \mathbb{R} \\ (F, G) & & \langle F, G \rangle_\Delta \end{array}$$

The same results occur with  $\nabla$  instead of  $\Delta$  and  $\iota_0^-$  instead of  $\iota_0^+$ .

In other word, the discrete scalar product is just the usual  $L^2$  scalar product with  $L^2$  functions replaces by discrete functions, the usual multiplication by its discrete analogue and the classical antiderivative by its discrete pendant. This phenomenon is in fact general. We will formalize this property in the next part using discrete embeddings formalisms.

## 6. Discrete classical results

In this Section, we derive two discrete analogue of classical results in Analysis : the fundamental theorem of differential calculus and the Dubois-Raymond lemma.

**6.1. Discrete fundamental theorem of differential calculus.** — As we are looking for the transfer of algebro-analytic properties of derivatives and antiderivatives a natural question is up to which extent the *fundamental theorem of differential calculus* (see [21], Theorem 6.13 p.239) is preserved ? A very nice feature of our derivation is that this theorem is the following discrete analogue of this result:

**Theorem 24 (Fundamental theorem of the discrete differential calculus)**

For all  $F \in C(\mathbb{T}, \mathbb{R})$  we have

$$(10) \quad \begin{aligned} J_\Delta \circ \Delta(F) &= F - \mathbb{F}_0 \quad \text{and} \quad \Delta \circ J_\Delta F = F, \\ J_\nabla \circ \nabla(F) &= F - \mathbb{F}_0 \quad \text{and} \quad \nabla \circ J_\nabla F = F. \end{aligned}$$

*Proof.* — We make the proof only for  $\Delta$  and  $J_\Delta$ . The proof for  $\nabla$  and  $J_\nabla$  being similar.

As  $d^+ \circ \iota_1(F) \in P_0^+([a, b], \mathbb{R})$ , we deduce that  $\iota_0^+ \circ \pi(d^+ \circ \iota_1(F)) = d^+ \circ \iota_1(F)$ . As a consequence, we obtain

$$(11) \quad \begin{aligned} J_\Delta \circ \Delta(F) &= \pi \circ \int_a^t \circ \iota_0^+ \circ \pi \circ d^+ \circ \iota_1(F) = \pi \circ \int_a^t d^+(\iota_1(F)), \\ &= \pi(\iota_1(F)(t) - \iota_1(F)(0)) = F - \mathbb{F}_0. \end{aligned}$$

Second we remark that over  $P_1([a, b], \mathbb{R})$  we have  $\iota_1 \circ \pi = Id$ . As  $\int_a^t \iota_0^+(F) \in P_1([a, b], \mathbb{R})$  then by definition we have

$$(12) \quad \begin{aligned} \Delta \circ J_\Delta(F) &= \pi \circ d^+ \circ \iota_1 \circ J_\Delta(F) = \pi \circ d^+ \circ \iota_1 \circ \pi \circ \int_a^t \iota_0^+(F), \\ &= \pi \left( d^+ \circ \int_a^t \iota_0^+(F) \right) = \pi(\iota_0^+(F)(t)) = F. \end{aligned}$$

This concludes the proof. □

**6.2. Discrete Dubois-Raymond lemma.** — The discrete version of the Dubois-Raymond lemma is valid. We first introduce the set  $C_0(\mathbb{T}, \mathbb{R}) \subset C(\mathbb{T}, \mathbb{R})$  defined by

$$(13) \quad C_0(\mathbb{T}, \mathbb{R}) = \{G \in C(\mathbb{T}, \mathbb{R}), G_0 = G_N = 0\}.$$

**Lemma 25.** — Let  $F \in C(\mathbb{T}, \mathbb{R})$  such  $[J_\Delta(F \star G)]_N = 0$  for all  $G \in C_0(\mathbb{T}, \mathbb{R})$  then  $F_k = 0$  for  $k = 1, N - 1$ .

*Proof.* — Let  $G \in C_0(\mathbb{T}, \mathbb{R})$ . We choose  $G$  such that  $G_k = F_k$  for  $1 \leq k \leq N-1$ . Hence, we obtain  $[J_\Delta(F \star G)]_N = \sum_{k=1}^{N-1} F_k^2 h = 0$ . We deduce  $F_k = 0$  for  $k = 1, \dots, N-1$ . This concludes the proof.  $\square$

A stronger result can be derived when  $F$  is replaced by  $\Delta F$  :

**Lemma 26.** — *Let  $F \in C(\mathbb{T}, \mathbb{R})$  such that  $[J_\Delta((\Delta F) \star G)]_N = 0$  for all  $G \in C(\mathbb{T}_+, \mathbb{R})$  then  $\Delta F = 0$ .*

*Proof.* — The proof follows from a simple computation. As  $[J_\Delta(\Delta F \star G)]_N = \sum_{k=0}^{N-1} \frac{F_{k+1} - F_k}{h} G_k h$ , we obtain, taking  $G = \Delta F \in C(\mathbb{T}_+, \mathbb{R})$  that  $F_{k+1} - F_k = 0$  for all  $k = 0, \dots, N-1$ , so that  $F = F_0$ . As a consequence,  $F$  is a constant for  $\Delta$  which concludes the proof.  $\square$

Similar results can be obtained for the  $\nabla$ -derivative.

## PART II

### DISCRETE EMBEDDING FORMALISMS

We use the previous abstract approach to finite differences in order to provide a formal definition of a discrete analogue for differential equations, functional and other objects which can be defined using integrals, derivatives and functions. We first define a general abstract procedure and give three natural discrete generalisation of a differential equations. This point of view allows us to more precisely determine on which assumptions a discretisation can be constructed.

#### 7. Abstract discrete embedding

In this Section, we consider a general *formal functional* made of symbols  $d/dt$  and  $\int$  acting on a given set of functions  $x, y$ , etc. We denote such a formal functional by

$$(14) \quad F(t, x, y; d/dt, \int),$$

as long as the expression defined by  $F$  is well defined.

A *formal relation* will be the data of a formal functional satisfying

$$(15) \quad F(t, x, y; d/dt, \int) = 0.$$

A classical formal relation is given by a first order differential equation  $dx/dt - f(x, t) = 0$ .

We can now define what is a discrete embedding of an abstract functional or relation.

**Definition 27 (Abstract discrete embedding).** — Let  $\mathcal{E}$  be a given discrete embedding and  $F$  be a formal functional defined by (14). The discrete embedding of  $F$  denoted by  $F_{\mathcal{E}}$  is defined for all  $X \in \mathcal{E}$  by

$$(16) \quad F_{\mathcal{E}}(T, X) = F(T, X; (d/dt)_{\mathcal{E}}, \int_{\mathcal{E}}),$$

where  $\mathcal{E}$ ,  $(d/dt)_{\mathcal{E}}$  and  $\int_{\mathcal{E}}$  are the discrete functional space, the discrete derivative and discrete antiderivative which are fixed.

**Remark 2.** — The previous procedure deals with the minimal objects used to write a given differential relation. In many situations as for example geometry, a differential relation put in evidence some particular operators, like the Laplacian, from which a differential equations is constructed. In such a context, one can be conducted to define a discrete embedding directly focusing on the given operator and its algebraic or geometric properties. The discrete embedding then follows the same lines as in Definition 27 but the "functorial" property that we are looking for is destroyed. We refer to [13] for more details.

## 8. Application to ordinary differential equations : the three forms

In this Section, we apply the previous formalism for ordinary differential equations. Using different representations of a given differential equation (differential or integral form, variational), we obtain discrete analogues which do not give always the same object. We have then multiplicity of *discrete realisations* of a given differential equations. This problem is in fact relevant in all embedding formalisms and leads to the *coherence problem* asking for conditions under which such representations coincide.

**8.1. Discrete differential embedding.** — Let  $x \in \mathbb{R}^d$ , we consider the ordinary differential equations

$$(17) \quad \frac{dx}{dt} = f(t, x).$$

Using the finite differences embedding, the discrete  $\Delta$  version of this equation is

$$(18) \quad \Delta X = f(T, X), X \in C(\mathbb{T}, \mathbb{R}^d), T \in C(\mathbb{T}, \mathbb{R})$$

As  $\Delta X$  is defined on  $\mathbb{T}_{\Delta}$ , we obtain for each  $i = 0, \dots, N-1$

$$(19) \quad \frac{X_{i+1} - X_i}{h} = f(T_i, X_i),$$

where  $T_i = a + i \frac{b-a}{h}$ .

Also we have the discrete  $\nabla$  version of this equation which is

$$(20) \quad \nabla X = f(T, X), X \in C(\mathbb{T}, \mathbb{R}^d), T \in C(\mathbb{T}, \mathbb{R}).$$

As  $\nabla X$  is defined on  $\mathbb{T}_{\nabla}$ , we obtain for each  $i = 1, \dots, N$

$$(21) \quad \frac{X_i - X_{i-1}}{h} = f(T_i, X_i),$$

**8.2. Discrete integral embedding.** — The integral formulation of the previous ordinary differential equation is given by

$$(22) \quad x(t) = x(0) + \int_0^t f(s, x(s)) ds.$$

The forward discrete embedding of this equation is then given by

$$(23) \quad X = \mathbb{X}_0 + J_\Delta(f(T, X)),$$

which we call the *forward discrete integral embedding*. By definition, this equation is equivalent to  $X_i = X_0 + h \sum_{j=0}^{i-1} f(T_j, X_j)$ ,  $i = 1, \dots, N$ . As a consequence, we obtain  $X_{i+1} - X_i = hf(T_i, X_i)$ ,  $i = 0, \dots, N-1$ , which is the classical *one step forward Euler scheme* in Numerical Analysis ([14], Chap.V, §.2.3).

In the same way one can obtain the *backward discrete integral embedding* of this equation which is then given by

$$(24) \quad X = \mathbb{X}_0 + J_\nabla(f(T, X)).$$

By definition, this equation is equivalent to  $X_i = X_0 + h \sum_{j=0}^{i-1} f(T_j, X_j)$ ,  $i = 1, \dots, N$ . Easy computations lead to  $X_i - X_{i-1} = hf(T_i, X_i)$ ,  $i = 0, \dots, N-1$ , which is the classical *one step backward Euler scheme* in Numerical Analysis.

**Remark 3.** — *The finite differences differential embedding of the equation is equivalent to the differential case. As a consequence, we see that in this simple case, we have coherence between the two discrete versions of the equation.*

**8.3. Discrete variational embedding.** — In this Section, we define the discrete variational embedding of a second order differential equation which is Lagrangian. Our derivation is compared with the classical work of J.E. Marsden and M. West [25] (see also [20]) about the discrete calculus of variation and variational integrators, in order to explain the interest of our abstract framework.

**8.3.1. Lagrangian systems.** — In this section, we recall classical definitions concerning Lagrangian systems.

**Definition 28.** — *A Lagrangian functional is an application denoted by  $\mathcal{L} : C^2([a, b], \mathbb{R}^d) \rightarrow \mathbb{R}$  and defined by  $\mathcal{L}q = \int_a^b L(q(t), \dot{q}(t), t) dt$ , where  $L : [a, b] \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  called Lagrangian and denoted by  $L(t, x, v)$ .*

The functional  $\mathcal{L}$  is also called action functional.

Let  $\mathcal{L}$  be a Lagrangian functional, we denote by  $D\mathcal{L}(q)(w)$  the Fréchet derivative of  $\mathcal{L}$  in  $q$  along the direction  $w$  in  $C^2([a, b], \mathbb{R}^d)$ , i.e.

$$(25) \quad D\mathcal{L}(q)(w) = \lim_{\epsilon \rightarrow 0} \frac{\mathcal{L}(q + \epsilon w) - \mathcal{L}(q)}{\epsilon}.$$

An *extremal* (or *critical point*) of a Lagrangian functional  $\mathcal{L}$  is a trajectory  $q$  such that  $D\mathcal{L}(q)(w) = 0$  for any *variations*  $w$  (i.e.  $w \in C_0^1([a, b], \mathbb{R}^d) := \{w \in C^1([a, b], \mathbb{R}^d), w(a) = w(b) = 0\}$ ).

**8.3.2. Discrete Lagrangian functionals.** — The discrete calculus of variations is defined over discrete Lagrangian functionals which are obtained using the discrete embedding that we have fixed. We introduce  $\mathbb{T}^\pm := \mathbb{T} \setminus \{t_0, t_N\}$ .

**Definition 29.** — Let  $L$  be an admissible Lagrangian function and  $\mathcal{L}$  the associated functional. The discrete forward Lagrangian functional  $\mathcal{L}_\Delta$  associated to  $L$  with the  $\Delta$  integral is defined by

$$(26) \quad \mathcal{L}_\Delta(X) = [J_\Delta(L(T, X, \Delta X))]_N.$$

The previous form is completely fixed once one has given a discrete embedding formalism. We clearly see the relation between the classical function and the discrete one.

**8.3.3. Comparison with Marsden-West definition.** — In ([25] p.363 and §.1.3 p.370-371), the authors define discrete Lagrangian system for which they do not preserve the classical form of the continuous functional. They introduce a discrete Lagrangian given by

$$(27) \quad L_d(X_k, X_{k+1}, h) = hL(X_k, \Delta(X)_k, t_k),$$

which gives the discrete functional

$$(28) \quad \mathcal{L}_h(X) = \sum_{k=0}^{N-1} L_d(X_k, X_{k+1}, h).$$

The algebraic structure of the classical functional is then destroyed. This point has an important consequence in the derivation of the discrete Euler-Lagrange equation in ([25], p.363 and §.1.3, Theorem 1.3.1 p.371). Indeed, as we will see, the form obtained in [25] does not put in evidence the complete analogy between the discrete Euler-Lagrange equation and the classical one.

The same remark applies to the presentation made by E. Hairer, C. Lubich and G. Wanner in the book *geometric numerical integration* (see [20], formula (6.5) and (6.6) p.192-195).

**8.3.4. Discrete calculus of variations and discrete Euler-Lagrange equations.** — An element of  $C_0(\mathbb{T}, \mathbb{R})$  is called a *discrete variation*. As  $C(\mathbb{T}, \mathbb{R})$  is a linear space, we can define the Frechet derivative of  $\mathcal{L}_\Delta$  (resp.  $\mathcal{L}_\nabla$ ) along a given direction  $H \in C(\mathbb{T}, \mathbb{R})$  and denoted by

$$(29) \quad \mathcal{DL}_\Delta(X)(H) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (\mathcal{L}_\Delta(X + \epsilon H) - \mathcal{L}_\Delta(X)).$$

The corresponding notion of critical points is given by:

**Definition 30.** — A discrete critical point  $X \in C(\mathbb{T}, \mathbb{R})$  verify  $\mathcal{DL}_\Delta(X)(H) = 0$  for all  $H \in C_0(\mathbb{T}, \mathbb{R})$ .

We obtain the following *discrete Euler-Lagrange equation* :

**Theorem 31 (Discrete Euler-Lagrange equation).** — Let  $L$  be an admissible Lagrangian function. A discrete function  $X \in C(\mathbb{T}, \mathbb{R})$  is a critical point of the discrete  $\Delta$  Lagrangian functional associated to  $L$  if and only if it satisfies

$$(30) \quad \nabla \left( \frac{\partial L}{\partial v}(T, X, \Delta X) \right) = \frac{\partial L}{\partial x}(T, X, \Delta X),$$

over  $\mathbb{T}^\pm$ .

The previous formulation makes clear the relation between the usual Euler-Lagrange equation and the discrete one. In particular, one see the duality between the two operators  $\Delta$  and  $\nabla$  which induces a natural mixing between the two discrete derivatives which is hidden in the continuous case.

*8.3.5. Comparison with Marsden-West definition.* — One can compare this writing of the discrete Euler-Lagrange equation with the one obtained in ([25], p.363). Using the form (28), they have

$$(31) \quad D_2 L_d(X_{k-1}, X_k, h) + D_1 L_d(X_k, X_{k+1}, h) = 0,$$

for  $k = 1, \dots, N-1$ . The usual form of the Euler-Lagrange equation is completely lost and by the way the analogy between the two objects.

The same remark applies to the presentation made by E. Hairer, C. Lubich and G. Wanner in the book *geometric numerical integration* (see [20], formula (6.7) p.192-195).

*8.3.6. Proof of Theorem 31.* — The proof can be done in different ways. However, we want to keep a proof which is similar to the classical proof of the continuous Euler-Lagrange equation as exposed for example in ([19], Theorem 1, p.15).

Using a Taylor expansion of  $L$ , we obtain

$$(32) \quad \mathcal{DL}_\Delta(X)(H) = \left[ J_\Delta \left( \frac{\partial L}{\partial v}(T, X, \Delta X) \star \Delta H + \frac{\partial L}{\partial x}(T, X, \Delta X) \star H \right) \right]_N.$$

As  $H \in C_0(\mathbb{T}, \mathbb{R})$ , we have using the discrete integration by parts formula

$$(33) \quad \mathcal{DL}_\Delta(X)(H) = \left[ J_\Delta \left( -\nabla \left( \frac{\partial L}{\partial v}(T, X, \Delta X) \right) \star H + \frac{\partial L}{\partial x}(T, X, \Delta X) \star H \right) \right]_N.$$

Using the discrete Dubois-Raymond lemma, we deduce

$$(34) \quad -\nabla \left[ \frac{\partial L}{\partial v}(T, X, \Delta X) \right] + \frac{\partial L}{\partial x}(T, X, \Delta X) = 0 \quad \text{over } \mathbb{T}^\pm.$$

This concludes the proof.

## PART III

### HIGHER ORDER DISCRETE EMBEDDINGS

As it is well-known there exist other way to obtain better approximation of discrete integral for numerical computations such as the midpoint method, the trapezoidal method and the Simpson method. Our choice of discrete integral correspond to the left method for the  $\Delta$  integral and the right method for the  $\nabla$  integral. In our construction of finite difference embedding, the choice of discrete integral is explicitly dependent of the choice of the discretisation of functions and the discrete derivative in order to preserve the fundamental theorem of differential calculus. In what follow we briefly present two higher interpolation method which leads to the midpoint method, the trapezoidal method and the Simpson method for the forward derivative only. The backward derivative will be straightforward obtained.



### 9. Quadratic embedding : Trapezoidal and Midpoint methods

In this Section, we obtain the classical one step Trapezoidal (resp. Midpoint) numerical scheme (see [14], III.1.2) as respectively the differential (resp. integral) discrete embedding of the quadratic discrete embedding that we define in the next Section.

**9.1. Quadratic discrete embedding.** — We follow the general strategy of discrete embeddings. We define a quadratic interpolation map which is used to obtain a discrete derivative denoted by  $\Delta_2$ . We then construct the corresponding discrete integral denoted by  $J_{\Delta_2}$ .

*9.1.1. Quadratic interpolation mapping.* — We suppose  $N = 2d, d \geq 1, d \in \mathbb{N}$ . We define a *quadratic interpolation mapping* denoted by  $\iota_2 : C(\mathbb{T}, \mathbb{R}) \rightarrow P_2([a, b], \mathbb{R})$  using Lagrange polynomial of degree 2. Precisely, we have

$$(35) \quad \iota_2(F)(t) = \sum_{i=0}^{d-1} P_{L,2}^{F_{2i}, F_{2i+1}, F_{2i+2}}(t) 1_{[t_{2i}, t_{2i+2}[}.$$

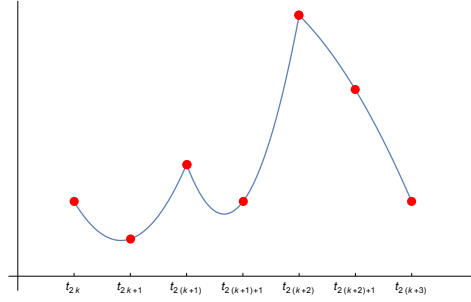


FIGURE 6. Quadratic interpolation mapping

*9.1.2. Quadratic discrete derivative and integral.* — We follow the strategy used in Part I. We denote by  $\Delta_2$  the discrete derivative defined by

$$(36) \quad \Delta_2 = \pi \circ d^+ \circ \iota_2.$$

An explicit form is given by

$$(37) \quad \begin{aligned} [\Delta_2(F)]_{2k} &= \frac{2(F_{2k+1} - F_{2k})}{h} - \frac{(F_{2k+2} - F_{2k})}{2h}, \\ [\Delta_2(F)]_{2k+1} &= \frac{(F_{2k+2} - F_{2k})}{2h}, \end{aligned}$$

for  $k = 0, \dots, d-1$ .

In order to define the discrete integral associated to  $\Delta_2$ , we must find an interpolation mapping denoted by  $\iota_1^+$  in the following, such that  $\iota_1^+ \circ \pi = \text{Id}$  on  $P_1^+([a, b])$ . We have

$$(38) \quad \iota_1^+(F)(t) = \sum_{k=0}^{d-1} \left[ F_{2k} + \frac{F_{2k+1} - F_{2k}}{h} (t - t_{2k}) \right] 1_{[t_{2k}, t_{2k+2}[}(t).$$

The discrete integral is then defined by

$$(39) \quad J_{\Delta_2} = \pi \circ \int_a^t \circ \iota_1^+,$$

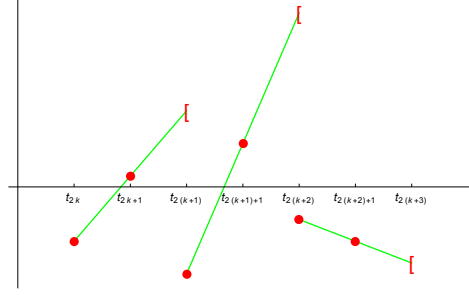


FIGURE 7. Quadratic forward derivative

and satisfies  $J_{\Delta_2}(\Delta_2(F)) = F - \mathbb{F}_0$ . Indeed, we have  $J_{\Delta_2}(\Delta_2(F)) = \pi \circ \int_a^t \circ \tau_1^+ \circ \pi \circ d^+ \circ \iota_2(F)$ . As  $d^+ \circ \iota_2(F) \in P_1^+([a, b])$  and  $\tau_1^+ \circ \pi = \text{rm Id}$  on  $P_1^+([a, b])$ , this equality reduces to  $J_{\Delta_2}(\Delta_2(F)) = \pi \circ \int_a^t \circ d^+ \circ \iota_2(F) = \pi [\iota_2(F) - F_0] = F - \mathbb{F}_0$ .

An explicit form for  $J_{\Delta_2}$  is given by

$$(40) \quad [J_{\Delta}(F)]_{2k} = \sum_{q=0}^{k-1} 2hF_{2q+1},$$

for  $k = 1, \dots, d$  with  $[J_{\Delta_2}(F)]_0 = 0$  and

$$(41) \quad [J_{\Delta_2}(F)]_{2k+1} = h \frac{F_{2k} + F_{2k+1}}{2} + \sum_{q=0}^{k-1} 2hF_{2q+1},$$

for  $k = 1, \dots, d-1$  with  $[J_{\Delta_2}(F)]_1 = h \frac{F_0 + F_1}{2}$ .

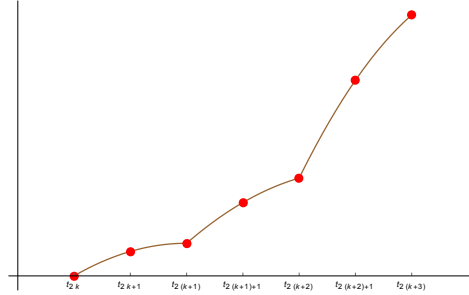


FIGURE 8. Quadratic discrete integral

**9.2. Discrete differential and integral embedding.** — Applying the previous discrete embedding, we obtain the following result :

**Theorem 32.** — *The  $\Delta_2$  differential embedding (resp. integral embedding) correspond to the one step Trapezoidal (resp. Midpoint) scheme.*

A direct consequence is the following corollary :

**Corollary 33.** — *The  $\Delta_2$  discrete embedding is not coherent.*

It is interesting to notice that the  $\Delta$  discrete embedding was coherent with respect to the differential/integral embedding only breaking when one considers variational discrete embedding. In this case, non-coherence already break-up at the first level.

## 10. Cubic discrete embedding : the forward Simpson numerical scheme

In this Section, we consider a cubic discrete embedding. The main point is that we recover coherence between the differential/integral embedding and obtain the *forward Simpson numerical scheme* (see [14], III.1.2).

### 10.1. Cubic discrete embedding. —

*10.1.1. Cubic interpolation mapping.* — We suppose  $N = 3d, d \geq 1, d \in \mathbb{N}$ . We define a *cubic interpolation mapping* denoted by  $\iota_3 : C(\mathbb{T}, \mathbb{R}) \rightarrow P_3([a, b], \mathbb{R})$  using Lagrange polynomial of degree 3. Precisely, we have

$$(42) \quad \iota_3(F)(t) = \sum_{i=0}^{d-1} P_{L,3}^{F_{3i}, F_{3i+1}, F_{3i+2}, F_{3i+3}}(t) 1_{[t_{3i}, t_{3i+3}[}.$$

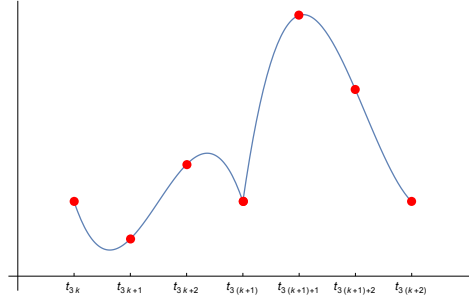


FIGURE 9. Cubic interpolation mapping

*10.1.2. Cubic discrete derivative and integral.* — We follow the strategy used in Part I. We denote by  $\Delta_3$  the discrete derivative defined by

$$(43) \quad \Delta_3 = \pi \circ d^+ \circ \iota_3.$$

An explicit form for  $\Delta_3$  is given by

$$(44) \quad \begin{aligned} [\Delta_3(F)]_{3k} &= -3 \frac{(F_{3k+2} - 2F_{3k+1} + F_{3k})}{2h} - \frac{(F_{3k+3} - F_{3k})}{3h}, \\ [\Delta_3(F)]_{3k+1} &= -\frac{(F_{3k+1} - 2F_{3k+2} + F_{3k})}{2h} - \frac{(F_{3k+3} - F_{3k})}{6h}, \\ [\Delta_3(F)]_{3k+2} &= \frac{(F_{3k+3} - 2F_{3k+1} + F_{3k+2})}{2h} - \frac{(F_{3k+3} - F_{3k})}{6h}, \end{aligned}$$

for  $k = 0, \dots, d-1$ .

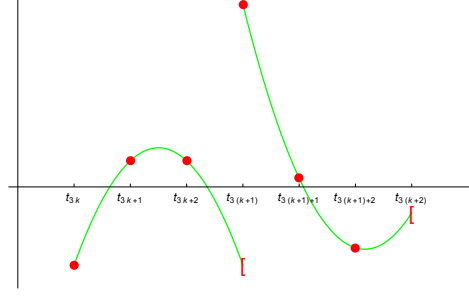


FIGURE 10. Cubic forward derivative

In order to define the discrete integral associated to  $\Delta_3$ , we must find an interpolation mapping denoted by  $\iota_2^+$  in the following, such that  $\iota_2^+ \circ \pi = \text{Id}$  on  $P_2^+([a, b])$ . We have

$$(45) \quad \iota_2^+(F)(t) = \sum_{k=0}^{d-1} \left[ \frac{F_{3k}(t - t_{3k+1})(t - t_{3k+2})}{(t_{3k} - t_{3k+1})(t_{3k} - t_{3k+2})} - \frac{F_{3k+1}(t - t_{3k})(t - t_{3k+2})}{(t_{3k} - t_{3k+1})(t_{3k+1} - t_{3k+2})} + \frac{F_{3k+2}(t - t_{3k})(t - t_{3k+1})}{(t_{3k+2} - t_{3k})(t_{3k+2} - t_{3k+1})} \right] 1_{[t_{3k}, t_{3k+3}]}(t).$$

The discrete integral is then defined by

$$(46) \quad J_{\Delta_3} = \pi \circ \int_a^t \circ \iota_2^+,$$

and satisfies  $J_{\Delta_3}(\Delta_3(F)) = F - F_0$ . Indeed, we have  $J_{\Delta_3}(\Delta_3(F)) = \pi \circ \int_a^t \circ \iota_2^+ \circ \pi \circ d^+ \circ \iota_3(F)$ . As  $d^+ \circ \iota_3(F) \in P_2^+([a, b])$  and  $\iota_1^+ \circ \pi = \text{rm Id}$  on  $P_2^+([a, b])$ , this equality reduces to  $J_{\Delta_2}(\Delta_2(F)) = \pi \circ \int_a^t \circ d^+ \circ \iota_3(F) = \pi [\iota_3^+(F) - F_0] = F - F_0$ .

An explicit form for  $J_{\Delta_3}$  is given by

$$(47) \quad [J_{\Delta_3}(F)]_{3k} = \sum_{q=0}^{k-1} 3h \frac{(F_{3q} + 3F_{3q+2})}{4},$$

for  $k = 1, \dots, d$  and by

$$[J_{\Delta_3}(F)]_{3k+1} = \sum_{q=0}^{k-1} \frac{h}{12} (5F_{3q} + 8F_{3q+1} - F_{3q+2})$$

$$[J_{\Delta_3}(F)]_{3k+2} = \sum_{q=0}^{k-1} \frac{h}{3} (F_{3q} + 4F_{3q+1} + F_{3q+2}),$$

for  $k = 1, \dots, d-1$ .

**10.2. Discrete differential and integral embedding.** — We have the following Theorem :

**Theorem 34.** — *The  $\Delta_3$  differential and integral discrete embedding corresponds to the one forward Simpson numerical scheme.*

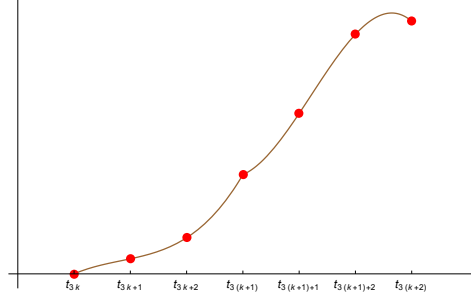


FIGURE 11. Cubic discrete integral

A direct consequence of this Theorem is :

**Corollary 35.** — *The  $\Delta_3$  discrete embedding is coherent.*

## 11. Conclusion and perspectives

The previous formalism includes the classical numerical methods in a unified point of view called *discrete embedding formalisms* which is itself an illustrative example of the philosophy of *embedding formalisms*. Doing so, we obtain others formulations of classical results like the discrete calculus of variations, the discrete Euler-Lagrange equation, discrete integration by part, etc which allows us to clearly put in evidence the correspondence between the continuous case and the discrete case. As explained in the last Section, this procedure is not limited to first order methods but can be used for arbitrary high order. Moreover, the strategy can also be adapted in order to consider some particular equations based on very specific differential operators. An interesting problem with respect to our approach is to give a full categorical formulation of discrete embedding formalisms.

## A

### Proof of Theorem 32

The proof is based on explicit computations of the different quantities.

The discrete  $\Delta_2$  differential embedding of (17) is given by  $\Delta_2 X = f(T, X)$  which reduces to

$$\begin{aligned} \frac{2(X_{2k+1} - X_{2k})}{h} - \frac{(X_{2k+2} - X_{2k})}{2h} &= f(T_{2k}, X_{2k}), \\ \frac{(X_{2k+2} - X_{2k})}{2h} &= f(T_{2k+1}, X_{2k+1}), \end{aligned}$$

for  $k = 0, \dots, d-1$ .

The discrete  $\Delta_2$  integral embedding of (22) is given by  $X = X_0 + J_{\Delta_2}(f(T, X))$  which gives

$$(48) \quad X_{2k} = X_0 + \sum_{q=0}^{k-1} 2hf(T_{2q+1}, X_{2q+1}),$$

for  $k = 1, \dots, d$  and

$$(49) \quad X_{2k+1} = X_0 + h \frac{f(T_{2k}, X_{2k}) + f(T_{2k+1}, X_{2k+1})}{2} + 2h \sum_{q=0}^{k-1} f(T_{2q+1}, X_{2q+1}),$$

for  $k = 0, \dots, d-1$ .

As a consequence, the discrete  $\Delta_2$ -differential embedding is equivalent to

$$X_{2k+1} = X_{2k} + h \frac{f(T_{2k}, X_{2k}) + f(T_{2k+1}, X_{2k+1})}{2}, \text{ for } k = 0, \dots, d-1,$$

which is the classical one step *trapezoidal scheme* in Numerical Analysis (see [14], III.1.2 (c) p.63) and also

$$X_{2k+2} = X_{2k} + 2h f(T_{2k+1}, X_{2k+1})$$

which is the classical one step *forward midpoint scheme* in Numerical Analysis (see [14], III.1.2 (a) p.61). This concludes the proof.

## B

### Proof of Theorem 34

The proof of Theorem 34 is based on explicit computations. The discrete  $\Delta_3$  differential embedding of (17) is given by  $\Delta_3 X = f(T, X)$  which reduces to

$$\begin{aligned} -3 \frac{(F_{3k+2} - 2F_{3k+1} + F_{3k})}{2h} - \frac{(F_{3k+3} - F_{3k})}{3h} &= f(T_{3k}, X_{3k}), \\ -\frac{(F_{3k+1} - 2F_{3k+2} + F_{3k})}{2h} - \frac{(F_{3k+3} - F_{3k})}{6h} &= f(T_{3k+1}, X_{3k+1}), \\ \frac{(F_{3k+3} - 2F_{3k+1} + F_{3k+2})}{2h} - \frac{(F_{3k+3} - F_{3k})}{6h} &= f(T_{3k+2}, X_{3k+2}), \end{aligned}$$

for  $k = 0, \dots, d-1$ .

The discrete  $\Delta_3$  integral embedding of (22) is given by  $X = \mathbb{X}_0 + J_{\Delta_3}(f(T, X))$  which gives

$$(50) \quad X_{3k} = X_0 + \sum_{q=0}^{k-1} 3h \frac{(f(T_{3q}, X_{3q}) + 3f(T_{3q+2}, X_{3q+2}))}{4},$$

for  $k = 1, \dots, d$  and

$$\begin{aligned} X_{3k+1} &= X_0 + \sum_{q=0}^{k-1} \frac{h}{12} (5f(T_{3q}, X_{3q}) + 8f(T_{3q+1}, X_{3q+1}) - f(T_{3q+2}, X_{3q+2})), \\ X_{3k+2} &= X_0 + \sum_{q=0}^{k-1} \frac{h}{3} (f(T_{3q}, X_{3q}) + 4f(T_{3q+1}, X_{3q+1}) + f(T_{3q+2}, X_{3q+2})), \end{aligned}$$

for  $k = 0, \dots, d-1$ .

As a consequence, we obtain for  $k = 0, \dots, d - 1$

$$X_{3k+2} = X_{3k} + \frac{h}{3} (f(T_{3k}, X_{3k}) + 4f(T_{3k+1}, X_{3k+1}) + f(T_{3k+2}, X_{3k+2})),$$

which is the classical *forward Simpson scheme* in Numerical Analysis (see [14], III.1.2 (c) p.63). This concludes the proof.

## References

- [1] I.D. Albu and D. Oprea. Helmholtz type condition for mechanical integrators. *Novi Sad J. Math.*, 29(3):11–21, 1999. XII Yugoslav Geometric Seminar (Novi Sad, 1998).
- [2] V. I. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1979.
- [3] S. Awodey, *Category Theory*, Oxford Logic Guides Vol. 52, 2d Edition, Oxford University Press, 2011.
- [4] L. Bourdin, Nonshifted calculus of variations on time scales with Nabla-differentiable Sigma, *Journal of Mathematical Analysis and Applications*, 411(2):543-554, 2014.
- [5] L. Bourdin, *Contributions au calcul des variations et au principe du maximum de Pontryagin en calculs time scale et fractionnaire*, Ph.D. Thesis, Université de Pau et des Pays de l'Adour, 2013.
- [6] L. Bourdin, J. Cresson. Helmholtz's inverse problem of the discrete calculus of variations. *Journal of Difference Equations and Applications*, 19(9):1417–1436, 2013.
- [7] L. Bourdin, J. Cresson, I. Greff, A continuous/discrete fractional Noether's theorem, *Communication in Nonlinear Sciences and Numerical Simulations*, 18(4):878-887, 2013.
- [8] J. Cresson, Introduction to embedding of Lagrangian systems, *International Journal of Biomathematics and Biostatistics*, Vol. 1, no. 1, 23-31, 2010.
- [9] J. Cresson, Fractional variational embedding and Lagrangian formulations of dissipative partial differential equations, in *Fractional Calculus in Analysis, dynamics and optimal control*, p.65-127, Nova Publishers, New-York, 2013.
- [10] J. Cresson, S. Darses, Stochastic embedding of dynamical systems, *J. Math. Phys.* 48, 072703 (2007) (54 pages).
- [11] J. Cresson, I. Greff, C. Pierre, Coherent discrete embeddings for Lagrangian and Hamiltonian systems, 23.p, 2011.
- [12] J. Cresson, D. Torres, A. Malinowska, Differential, integral and variational time-scales embedding of Lagrangian systems, *Computers and Mathematics with Applications* Volume 64, Issue 7, October 2012, Pages 2290-2301.
- [13] J. Cresson, L. Hume, About the discrete Laplace operator, 2014.
- [14] J-P. Demailly, *Analyse numérique des équations différentielles*, Coll. Grenoble Sciences, 2006.
- [15] D. Crăciun and D. Oprea. The Helmholtz conditions for the difference equations systems. *Balkan J. Geom. Appl.*, 1(2):21–30, 1996.
- [16] D.R. Davis. The inverse problem of the calculus of variations in higher space. *Trans. Amer. Math. Soc.*, 30(4):710–736, 1928.
- [17] D.R. Davis. The inverse problem of the calculus of variations in a space of  $(n + 1)$  dimensions. *Bull. Amer. Math. Soc.*, 35(3):371–380, 1929.
- [18] J. Douglas. Solution of the inverse problem of the calculus of variations. *Trans. Amer. Math. Soc.*, 50:71–128, 1941.
- [19] I.M. Gelfand, S.V. Fomin, *Calculus of variations*, Dover ed., 2000.
- [20] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006. Structure-preserving algorithms for ordinary differential equations.

- [21] E. Hairer, G. Wanner, *Analysis by its history*, Undergraduate Texts in Mathematics, Spriner, 1996.
- [22] H. Helmholtz. Über der physikalische bedeutung des princips der kleinsten wirkung. *J. Reine Angew. Math.*, 100:137–166, 1887.
- [23] A. Hirsch. Über eine charakterische eigenschaft der differentialgleichungen der variationsrechnung. *Math. Ann.*, 49:49–72, 1897.
- [24] A. Hirsch. Die existenzbedingungen des verallgemeinsterten kinetischen potentialen. *Math. Ann.*, 50:429–441, 1898.
- [25] J.E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numer.*, 10:357–514, 2001.
- [26] A. Mayer. Die existenzbedingungen eines kinetischen potentiales. *Math-Phys. Kl.*, 84:519–529, 1896.
- [27] P.J. Olver. *Applications of Lie groups to differential equations*, volume 107 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1993.
- [28] *Foundations of theoretical mechanics*. Foundations of Theoretical Mechanics. Springer-Verlag, 1978.
- [29] V. Volterra. *Leçons sur les fonctions de lignes*. Gauthier-Villars. Paris, 1913.

---

JACKY CRESSON<sup>1,2</sup> AND FRÉDÉRIC PIERRET<sup>1</sup>, <sup>1</sup> SYRTE, Observatoire de Paris, 77 avenue Denfert-Rochereau, 75014 Paris, France